HHS Public Access

Author manuscript

J Neurosci Methods. Author manuscript; available in PMC 2018 April 15.

Published in final edited form as:

J Neurosci Methods. 2017 April 15; 282: 1–8. doi:10.1016/j.jneumeth.2017.02.009.

Automatic detection of periods of slow wave sleep based on intracranial depth electrode recordings

Chrystal M. Reed[1], **Kurtis G. Birch**[2], **Jan Kamiński**[2], **Shannon Sullivan**[2], **Jeffrey M. Chung**[1], **Adam N. Mamelak**[2], and **Ueli Rutishauser**[1,2]

[1]Department of Neurology, Cedars-Sinai Medical Center, Los Angeles, California, USA

[2]Department of Neurosurgery, Cedars-Sinai Medical Center, Los Angeles, California, USA

Abstract

Background—An automated process for sleep staging based on intracranial EEG data alone is needed to facilitate research into the neural processes occurring during slow wave sleep (SWS). Current manual methods for sleep scoring require a full polysomnography (PSG) set-up, including electrooculography (EOG), electromyography (EMG), and scalp electroencephalography (EEG). This set-up can be technically difficult to place in the presence of intracranial EEG electrodes. There is thus a need for a method for sleep staging based on intracranial recordings alone.

New Method—Here we show a reliable automated method for the detection of periods of SWS solely based on intracranial EEG recordings. The method is utilizes the ratio of spectral power in delta, theta, and spindle frequencies relative to alpha and beta frequencies to classify 30-second segments as SWS or not.

Results—We evaluated this new method by comparing its performance against visually scored patients (n=9), in which we also recorded EOG and EMG simultaneously. Our method had a mean positive predictive value of 64% across all nights. Also, an ROC analysis of the performance of our algorithm compared to manually labeled nights revealed a mean average area under the curve of 0.91 across all nights.

Comparison with existing method—Our method had an average kappa score of 0.72 when compared to visual sleep scoring by an independent blinded sleep scorer.

Conclusion—This shows that this simple method is capable of differentiating between SWS and non-SWS epochs reliably based solely on intracranial EEG recordings.

Keywords

Automatic sleep staging; Slow wave sleep; Electroencephalography; Intracranial EEG; Vigilance Index

1. Introduction

Sleep is, among many other functions, important for the consolidation of memories[1,24]. Also, certain types of focal seizures have a propensity to occur in different stages of the sleep-wake cycle[13,31]. In addition, the presence of interictal discharges specifically during some sleep-stages but not others might have potential localizing value[20,22,27]. Patients with medically refractory epilepsy often require the insertion of invasive intracranial electrodes to better localize their seizure focus. This provides the opportunity to record intracranial activity during different phases of sleep, which provides an opportunity to test different hypothesis on the role of sleep in normal function as well as epilepsy. There is thus substantial need to sleep stage such patients. However, traditional visual sleep staging requires three types of signals: scalp EEG, EOG, and EMG. Also, visual sleep staging can be subject to large interobserver variation[29], and requires substantial amounts of time for a rater to become accurate[23]. The scalp EEG set-up required for visual sleep staging is technically challenging to place in the presence of the head-wrap accompanying intracranial electrodes. Additionally, there exists no guidelines to perform visual sleep staging using intracranial EEG recordings. Consequently, there is a need for an automated sleep-scoring technique that relies only on intracranial signals. The availability of such a method would make it possible to perform analysis of sleep-related activity in cases where a simultaneous scalp EEG, EOG, and EMG was not available for sleep staging.

A review of automated sleep staging methods used in the past report an agreement as high as 98% with visual sleep scoring[32]. However, most such methods require a full polysomnography (PSG) set-up. Attempts to use single-channel EEG based automated sleep staging in humans have achieved a sensitivity index of 75–88%[5,11,16,35], however none of these in humans used intracranial EEG (Table 2).

Researchers focused on memory consolidation processes are particularly interested in slow wave sleep (SWS)[25]. Based on the 2012 revised guidelines, SWS is now referred to as Stage N3 sleep[4], however for the purpose of this study we will continued to refer to this sleep stage as SWS.

In humans, the longest epochs of SWS occur in the early part of the night and progressive shorten (in exchange for increased REM sleep) as the night progresses. The electrophysiological hallmark of SWS is high amplitude slow wave activity (1–4 Hz), as well as hippocampal sharp wave ripple oscillations[25]. Notably, SWS is not unique to humans: it has been reported in mammals[37], birds[19], and reptiles[28]. Automated methods to exclusively identify SWS have been few[3,28,33], and the inter-observer agreement can be as low as 67.2%[23].

Here, we present a new method for the automated detection of points of time during which a person is in SWS from intracranial depth electrode recordings alone. Our method is based on the vigilance index (VI), which is a previously defined index[17] equal to the ratio of the mean spectral power in specific frequency ranges (see methods). Previous scalp-EEG work showed that values of VI were high during SWS and low during wakefulness[17]. Here, we test the hypothesis that the VI index can be used to determine the presence or absence of

SWS based on intracranial depth electrode recordings alone. We tested this hypothesis by comparing the VI-based algorithm's performance with that of human raters in nine patients. These raters used simultaneously recorded scalp EEG, EMG and EOG to manually determine whether a patient is in SWS or not. Our automated method showed excellent inter-rater reliability (Kappa = 0.72), an average area under curve (AUC) of 0.91, and an average 64% positive predictive value for our automatically set threshold when compared to manual sleep staging. We thus suggest that this simple method is a way to determine whether patients are in SWS or not based on intracranial EEG alone.

2. Methods

Sleep recordings

Nine epileptic patients (see Table 1) undergoing invasive EEG monitoring with depth electrodes to localize seizures were continuously recorded for a 10-hour period during the night. Consecutive recordings were obtained using a Grass Tele factor (3 patients, 400 Hz sampling rate) or Neuralynx Atlas (6 patients, 2000 Hz sampling rate) system. No patients were excluded from our study. The patients were videotaped continuously. Sleep recordings were completed prior to sleep deprivation periods that are used clinically for eliciting seizures. The nights selected for sleep recordings were early on in the hospital stay while the patient's antiepileptic drugs were similar to their home dosing. Thus their sleep architecture should be similar to baseline and unperturbed by the prolonged hospital stay. This work was approved by the institutional review board of the Cedars-Sinai Medical Center, and all patients provided informed consent to participate. (IRB#13369).

Manual Sleep staging

In addition to intracranial recordings, we also recorded (only for the selected night) an electro-oculogram (EOG), electro-myogram (EMG), and 4 scalp EEG electrodes attached in standard locations (central channel referenced to an ear mastoid bilaterally) for polysomnography. Using the 2012 AASM sleep scoring manual definitions the entire night was visually sleep scored in consecutive 30 second epochs. Arousals were scored using the 2012 AASM sleep scoring manual definitions. For two patients, manual sleep staging was performed by two independent blinded scorers; User#2 was a board-certified clinical neurophysiologist.

The Vigilance index (VI)

Based on a postoperative MRI the 10 most lateral electrode contacts that were within brain tissue were selected for analysis (Figure 1, panels A, D, G). The brain regions consistently sampled for each patient include lateral temporal and lateral frontal lobe. In order to determine when a patient was in SWS, we first computed power spectra from 1 to 250 Hz (200 Hz for the patients 3 patients recorded on the Grass Telefactor) in different sleep stages. A pronounced difference was identified during SWS in the 1–4 Hz delta frequency band (Figure 2, Box). Our specific focus was to identify periods of such robust power difference in the delta frequency to identify SWS periods.

Unfiltered data from all 10 channels was processed in consecutive 30 second long non-overlapping epochs. The spectral power for each channel was first calculated using a fast Fourier transform (FFT). This was estimated for each segment and channel separately. We then normalized the total power of each channel by integrating the total area under the power spectrum and normalizing it to 1, so that channels with higher amplitude do not contribute more signal than those with low amplitudes. This normalized power spectrum was then averaged across all channels for each 30 second epoch. The delta, theta, spindle, alpha and beta power was then calculated for each 30s epoch. As previously reported[17], we defined VI as: VI= [delta power + theta power + spindle power]/[alpha power + high-beta power]. As frequency ranges we used delta (1–4 Hz), theta (4–7 Hz), spindle (11–16Hz), alpha (8–13 Hz) and high-beta (20–40 Hz)[17]. Low beta (14–20 Hz) was avoided since it overlaps with spindle frequencies. Each epoch will be associated with one VI value. After obtaining all VI values for all epochs of a night, we re-scaled all VI values for all the channels of that night such that the maximal VI value achieved was equal to 100 and the minimum to 1. This procedure makes VI values comparable between patients, but is otherwise not required (see results). All procedures were implanted in Matlab (Mathworks, Inc.).

Cohen's kappa coefficient

The Cohen's kappa coefficient (κ) is a statistical measure of inter-rater reliability between two raters[7] and was calculated as following[30]:

$$\kappa = \frac{\text{observed agreement} - \text{chance agreement}}{1 - \text{chance agreement}} = \frac{P_o - P_c}{1 - P_c}$$

where P_o is the proportion of observed agreements and P_c is the proportion of agreements expected by chance[30]. Here, P_c=0.5, since our algorithm identifies two possibilities (SWS and non-SWS). Kappa is considered to be a more robust measure than simple percent agreement calculations because κ takes into account only the agreements that occur above chance[30]. Following the sleep staging literature, the interpretation of kappa coefficients is as follows; values of ≤0.00 indicate poor agreement; 0.00–0.20 indicate slight agreement; 0.21–0.40 indicate fair agreement; 0.41–0.60 indicate moderate agreement; 0.61–0.80 indicate substantial agreement; and more than 0.80 indicates excellent agreement[18].

Sleep scoring algorithm

We defined the normalized VI range to be between 1 to 100. The closer VI to 100, the larger the probability the patient will be in SWS. We tested two different approaches to define a single parameter (the threshold). This threshold determines the value of VI above which an epoch is considered as SWS. The two methods tested are: i) threshold set equal to the value of VI at which a false positive rate of 0.1 was achieved for each night, and ii) threshold set equal to the mean VI plus one standard deviation of the VI for each night. We calculated κ values for these two different threshold setting methods for all nights to evaluate which method works best (Table 1). When using the algorithm for nights for which no visual sleep staging is available, only the second method is possible. Here, we in addition tested method 1 to evaluate how well the second method works.

Statistical algorithm

We performed a receiver operating characteristic (ROC) analysis to evaluate the performance of our algorithm. For each possible VI threshold value (1–100), we estimated the proportion of epochs correctly and incorrectly classified as SWS. This is referred to as true positive (TP) and false positive (FP) rate in the following. The true positive rate was defined as the epochs correctly identified as SWS. The false positive rate was defined as the epochs incorrectly identified as SWS. Similarly, true negative (TN) and false negative (FN) rates were defined. The specificity (TN/TN+FP), sensitivity (TP/TP+FN), and positive predictive value (TP/TP+FP), were calculated as a function of the threshold. For comparisons, we set the threshold to that which results in a false positive rate of 0.1. The area under the curve (AUC) of the resulting ROC curve for each night over all thresholds was calculated. Significant differences in VI between sleep stages were evaluated using un-paired t-tests ($p<0.001$). The kappa values for each night were obtained by after determining the thresholds with two different methods as described above. We tested the statistical significance of a particular κ value using a bootstrap statistic. For this purpose, we scrambled the ground truth labels assigned to each 30s epoch randomly and repeated this process 10,000 times to estimate the null distribution and mean chance κ value (see Table 1, row “mean kappa chance”).

3. Results

Spectral analysis of intracranial electrodes

The first goal of this study was to validate the VI concept derived from scalp EEG work for the use with intracranial depth electrode recordings. Twelve nights of sleep from nine consecutive patients were analyzed. An example recording, consisting of ten lateral intracranial EEG electrode contacts as well as scalp EEG and EOG is shown in Figure 1. Spectral analysis of the lateral ten intracranial EEG contacts of the first 7 patients revealed that spectral power in the delta frequency range was increased during SWS (Box, Figure. 2; green line). In contrast, alpha and beta frequency power decreased during SWS (Figure 2). This is expected[16,17] and we used this as a basis for the sleep scoring algorithm.

Calculating the Vigilance Index (VI)

Based on the spectral analysis of the 10 lateral intracranial electrodes, the VI was calculated for different sleep stages and wakefulness. We found that, for all nights, the average VI values were higher during SWS compared to both non-SWS and wakefulness (see Figure 4 for statistics). A representation of one such night in patient #7 is shown in Figures 3 and 4. The increase in VI correlated with the SWS scored by the manual scorer (Figure 3). This suggests the higher VI, the more likely the patient should be in SWS (Figure 4). We next verified this prediction by comparing to a manual sleep score for the same epochs.

AUC for VI compared to manual sleep staging

In order to verify the validity of this method for detection of SWS a receiver operating characteristic curve (ROC) was computed for each night over all the thresholds, with manual sleep staging as the ground truth. We then summarized the algorithms performance by

computing the area under the curve (AUC) of the ROC curve for each night compared to manual sleep staging. A representative ROC curve for patient #7 is shown in Figure 5A. The average AUC for the 12 nights was 0.91±0.05 (Table 1). To further validate the performance of our approach we compared the visual and automated sleep staging between two raters for two patients, which was similar (AUC of 0.88 vs. 0.83 for rater 1 and 2, respectively, Figure 6).

Comparing the computer algorithm with manual sleep staging

Above results show that an ROC analysis of a VI-based classifier that differentiates between SWS and non-SWS states performs well. However, this leaves open the question of how to automatically choose an appropriate threshold, i.e. a single point along the ROC curve. We tested two different methods to choose such a threshold (see methods): i) threshold set such that a false positive rate of 0.1 results, and ii) threshold set equal to the mean VI for plus one standard deviation of the VI (for each night). The purpose of using this approach was to calibrate our threshold finding method using data for which manual scoring was performed. We found that the average kappa for the VI threshold set at a false positive rate (FP) of 0.1 was 0.73±0.09 (Figure 5B, Table 1). The average (across nights) of the sensitivity, specificity and positive predictive value for the threshold set equal to that which results in an FP of 0.1 was 0.83±0.05, 0.83±0.04 and 64±12.50% respectively (Table 1, Figure 5C). The average kappa for the VI threshold set to the mean VI for each night plus one standard deviation was 0.72±0.12 (Table 1). Thus, setting the threshold using the second method, which does not require manual sleep staging, resulted in substantial agreement with manual sleep staging. The empirical p-value obtained by bootstrap statistics for each night with the threshold set for mean VI for each night plus one standard deviation was consistently 0.0001 (B=10’000 runs, p=1/B if none of the bootstrap samples exceeded the observed value). Thus demonstrating that the kappa values obtained for each night by using our set threshold was statistically higher than chance and is an accurate reflection of the kappa values obtained from manual sleep staging of SWS.

Replacing manual sleep staging with VI

In order for our automated procedure to replace visual scoring, we computed the VI for patient #9 prior to visual staging. Separately, we performed blinded visual staging. We found that the kappa for this patient was 0.78, when the threshold was set for mean plus one standard deviation of the mean VI for each night (Table 1). This indicates substantial agreement between the our automatic procedure and manual visual scoring, according to the criteria of Landis and Koch[18]. The average kappa for all the patients when the threshold VI was set to the mean plus one standard deviation of the mean VI was 0.72±0.12 (Table 1).

Artifacts

Artifacts produced by excessive movement can be seen in depth electrodes. Since the VI index is a ratio of slower to faster frequencies, we did not notice a spurious increase in VI due to excessive EMG artifact. Normalization of each patient’s VI is performed so that VI values are comparable between patients. However, this normalization step relies on the max VI value achieved during a given night and thus assumes that the patient entered SWS at least once during the night. However, in the case of sleep disorders with complete lack of

SWS[14] (which are rare), this normalization would not be appropriate. In this case, instead the unnormalized (actual) measured values have to be used to calculate VI. This would not impact performance of our algorithm if the detection threshold is chosen dynamically as a function of the mean VI of a given night (see above).

Epoch removal in the event of a seizure

Since all our patients have intractable epilepsy, there is a concern that they might have a seizure during the nights in which we were performing automated sleep staging. The rhythmic delta and theta activity seen at seizure onset and post-seizure caused an increase in VI. Indeed, patient #9 had a seizure during the night of sleep staging. The AUC for the sleep staging with the seizure was 0.611 (data not shown), the sleep staging was repeated after removal of epochs containing the seizure and post- ictal slowing, the AUC improved to 0.81 (Table 1). As expected the occurrence of the seizure lowered the kappa scores by 20–50%, especially when setting the threshold to the peak of specificity (data not shown).

4. Discussion

Our study demonstrated that using an automated vigilance index (VI) for SWS calculated from data recorded from 10 lateral (cortical) intracranial electrodes can potentially streamline the process of identifying periods of SWS.

Traditional sleep staging depends on at least 3-channel scalp EEG, 2-channel EOG and EMG. This method has been shown to vary with observers due to the highly subjective nature of distinguishing certain features[9]. This has led to the development of automated systems for EEG analysis during different stages of sleep (Table 2). Most of the automated systems still rely on signal from EEG, EMG and EOG, but a few systems have tried using single-channel surface EEG[2,5,10,12,16]. The problem with automated sleep staging is that none of them show a 100% agreement with manual sleep staging. This has led to the general belief that automated sleep staging system cannot replace manual sleep staging. The majority of disagreement between manual and automated sleep staging occurs during the transition periods and REM sleep (Table 2). However, for many research questions the highest priority is to detect SWS because it is particularly important for consolidation of hippocampal memories[24,38] and nocturnal seizures[13,27]. Our focus in this work was thus exclusively on detection of SWS in intracranial EEG. SWS contains predominantly high amplitude 1–4 Hz delta frequency power, compared to faster beta and alpha frequency power. Previous work on automated spectral analysis of frequency ratios during this stage using multi-channel[39] or single-channel scalp EEG in humans[16] has shown that such scoring can have an accuracy as high as 0.97[16,17,39]. Our method shows that a similar approach works well when only relying on intracranial human recordings. Using our automated VI, we obtained excellent agreement with human raters as demonstrated by the average AUC of 0.91 and an average PPV of 64%. Also, inter-rater reliability was high and when compared to visual sleep staging we obtained an average kappa coefficient of 0.72.

Central to our approach is the setting of a single parameter: a threshold, which determines the minimal VI value that is considered as SWS. This value is different between patients, but can be determined automatically. We determined that setting the threshold equal to the mean

plus one standard deviation of VI across the entire night works well. When the manual sleep staging was compared to the computer algorithm the average kappa obtained was 0.72 (Table 1) which is a substantial likelihood of the patient being in SWS. The same threshold when used to identify epochs of non-SWS; including REM, transitions states and non-stageable epochs has a higher kappa of 0.86. This indicates that our algorithm is fairly conservative, because a VI index lower than the mean for that night plus one standard deviation has a higher likelihood of not being SWS. The kappa coefficient and AUC obtained by performing a smoothing algorithm on the VI was not much different from the un-averaged VI. The occurrence of seizure and post-ictal slowing as expected decreased the AUC, and in order to avoid erroneous calculation of post-ictal delta slowing as SWS these epochs should be removed.

Applicability and limitations

Our study also has several limitations; our mean positive predictive value was 64% across all nights, substantially lower than the >90% achievable by other methods of automated sleep staging. Our attempt was to use this automated method to detect epochs of SWS, however our method cannot be used to identify epochs of REM or transition periods between stages. Hence, this study is not an attempt to replace manual sleep staging. Rather, we describe a method that can be used to exclusively identify SWS in human intracranial recordings, especially in situations where manual sleep staging is not possible. The number of epochs of SWS identified using our algorithm was sufficient to quantify sharp-wave ripples in the epileptogenic hippocampus during SWS[6]. Our automation can also be run retrospectively when visual sleep staging was not done. This allows for all nights recorded with intracranial EEG to be easily added to datasets when studying SWS relevant material without additional setup for PSG. Our method eliminates the subjectivity component of SWS scoring and speeds up the process for detection of SWS without sacrificing accuracy. Studying SWS is an important task for human neurophysiologists and the easy identification of SWS will aid the advancement of research in this area.

Acknowledgments

We thank the staff of the Epilepsy Monitoring Unit at Cedars-Sinai Medical Center for invaluable assistance.

Funding

This work was supported by the Neurosurgery Research and Education Foundation (to K.B.) and NIMH (R01MH110831, to U.R.).

References

1. Barrett TR, Ekstrand BR. Effect of sleep on memory. 3. Controlling for time-of-day effects. J Exp Psychol. 1972; 96:321–327. [PubMed: 4345763]
2. Baumgart-Schmitt R, Herrmann WM, Eilers R, Bes F. On the use of neural network techniques to analyse sleep EEG data. First communication: application of evolutionary and genetic algorithms to reduce the feature space and to develop classification rules. Neuropsychobiology. 1997; 36:194–210. [PubMed: 9396019]
3. Benington JH, Kodali SK, Heller HC. Scoring transitions to REM sleep in rats based on the EEG phenomena of pre-REM sleep: an improved analysis of sleep structure. Sleep. 1994; 17:28–36. [PubMed: 8191200]

4. Berry, RB,BR., Gamaldo, CE., Harding, SM., Lloyd, RM., Marcus, CL., Vaughn, BV. The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications. Darien, IL: 2015.
5. Berthomier C, Drouot X, Herman-Stoica M, Berthomier P, Prado J, Bokar-Thire D, Benoit O, Mattout J, d'Ortho MP. Automatic analysis of single-channel sleep EEG: validation in healthy individuals. Sleep. 2007; 30:1587–1595. [PubMed: 18041491]
6. Birch, KG., Mamelak, AN., Ross, IB., Chung, JM., Rutishauser, U. Increased Ratio of Hippocampal Ripples During Slow Wave Sleep (SWS) in the Epileptogenic Temporal Lobe Compared to the Normal Temporal Lobe. American Epilepsy Society; Philadelphia: 2015.
7. Cohen J. A coefficient of agreement for nominal scales. Educ Psychol Meas. 1960; 20:37–46.
8. Crisler S, Morrissey MJ, Anch AM, Barnett DW. Sleep-stage scoring in the rat using a support vector machine. J Neurosci Methods. 2008; 168:524–534. [PubMed: 18093659]
9. Danker-Hopfe H, Kunz D, Gruber G, Klosch G, Lorenzo JL, Himanen SL, Kemp B, Penzel T, Roschke J, Dorn H, Schlogl A, Trenker E, Dorffner G. Interrater reliability between scorers from eight European sleep laboratories in subjects with different sleep disorders. J Sleep Res. 2004; 13:63–69. [PubMed: 14996037]
10. Ebrahimi F, Mikaeili M, Estrada E, Nazeran H. Automatic sleep stage classification based on EEG signals by using neural networks and wavelet packet coefficients. Conf Proc IEEE Eng Med Biol Soc. 2008; 2008:1151–1154. [PubMed: 19162868]
11. Flexer A, Gruber G, Dorffner G. A reliable probabilistic sleep stager based on a single EEG signal. Artif Intell Med. 2005; 33:199–207. [PubMed: 15811785]
12. Fraiwan L, Lweesy K, Khasawneh N, Fraiwan M, Wenz H, Dickhaus H. Time frequency analysis for automated sleep stage identification in fullterm and preterm neonates. J Med Syst. 2011; 35:693–702. [PubMed: 20703521]
13. Herman ST, Walczak TS, Bazil CW. Distribution of partial seizures during the sleep–wake cycle: differences by seizure onset site. Neurology. 2001; 56:1453–1459. [PubMed: 11402100]
14. Higashi K, Sakata Y, Hatano M, Abiko S, Ihara K, Katayama S, Wakuta Y, Okamura T, Ueda H, Zenke M, Aoki H. Epidemiological studies on patients with a persistent vegetative state. J Neurol Neurosurg Psychiatry. 1977; 40:876–885. [PubMed: 599364]
15. Kohtoh S, Taguchi Y, Matsumoto N, Wada M, Huang ZL, Urade Y. Algorithm for sleep scoring in experimental animals based on fast Fourier transform power spectrum analysis of the electroencephalogram. Sleep and Biological Rhythms. 2008; 6:163–171.
16. Koley B, Dey D. An ensemble system for automatic sleep stage classification using single channel EEG signal. Comput Biol Med. 2012; 42:1186–1195. [PubMed: 23102750]
17. Kouider S, Andrillon T, Barbosa LS, Goupil L, Bekinschtein TA. Inducing task-relevant responses to speech in the sleeping brain. Curr Biol. 2014; 24:2208–2214. [PubMed: 25220055]
18. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977; 33:159–174. [PubMed: 843571]
19. Libourel PA, Herrel A. Sleep in amphibians and reptiles: a review and a preliminary analysis of evolutionary patterns. Biol Rev Camb Philos Soc. :2015.
20. Malow BA, Lin X, Kushwaha R, Aldrich MS. Interictal spiking increases with sleep depth in temporal lobe epilepsy. Epilepsia. 1998; 39:1309–1316. [PubMed: 9860066]
21. Mamelak A, Quattrochi JJ, Hobson JA. A microcomputer-based system for automated EEG collection and scoring of behavioral state in cats. Brain Res Bull. 1988; 21:843–849. [PubMed: 3219615]
22. Menezes Cordeiro I, von Ellenrieder N, Zazubovits N, Dubeau F, Gotman J, Frauscher B. Sleep influences the intracerebral EEG pattern of focal cortical dysplasia. Epilepsy Res. 2015; 113:132–139. [PubMed: 25986200]
23. Nizam A, Chen S, Wong S. Best-case kappa scores calculated retrospectively from EEG report databases. J Clin Neurophysiol. 2013; 30:268–274. [PubMed: 23733091]
24. Plihal W, Born J. Effects of early and late nocturnal sleep on declarative and procedural memory. J Cogn Neurosci. 1997; 9:534–547. [PubMed: 23968216]
25. Rasch B, Born J. About sleep's role in memory. Physiol Rev. 2013; 93:681–766. [PubMed: 23589831]

26. Rytkonen KM, Zitting J, Porkka-Heiskanen T. Automated sleep scoring in rats and mice using the naive Bayes classifier. J Neurosci Methods. 2011; 202:60–64. [PubMed: 21884727]
27. Sammaritano M, Gigli GL, Gotman J. Interictal spiking during wakefulness and sleep and the localization of foci in temporal lobe epilepsy. Neurology. 1991; 41:290–297. [PubMed: 1992379]
28. Shein-Idelson M, Ondracek JM, Liaw HP, Reiter S, Laurent G. Slow waves, sharp waves, ripples, and REM in sleeping dragons. Science. 2016; 352:590–595. [PubMed: 27126045]
29. Silber MH, Ancoli-Israel S, Bonnet MH, Chokroverty S, Grigg-Damberger MM, Hirshkowitz M, Kapen S, Keenan SA, Kryger MH, Penzel T, Pressman MR, Iber C. The visual scoring of sleep in adults. J Clin Sleep Med. 2007; 3:121–131. [PubMed: 17557422]
30. Sim J, Wright CC. The kappa statistic in reliability studies: use, interpretation, and sample size requirements. Phys Ther. 2005; 85:257–268. [PubMed: 15733050]
31. Singh S, Shukla G, Goyal V, Srivastava AK, Singh MB, Vibha D, Behari M. Impact of sleep on the localizing value of video EEG in patients with refractory focal seizures -a prospective video-EEG with EOG and submental EMG study. Clin Neurophysiol. 2014; 125:2337–2343. [PubMed: 24856459]
32. Sinha RK. Artificial neural network and wavelet based automated detection of sleep spindles, REM sleep and wake states. J Med Syst. 2008; 32:291–299. [PubMed: 18619093]
33. Stanus E, Lacroix B, Kerkhofs M, Mendlewicz J. Automated sleep scoring: a comparative reliability study of two algorithms. Electroencephalogr Clin Neurophysiol. 1987; 66:448–456. [PubMed: 2435525]
34. Stephenson R, Caron AM, Cassel DB, Kostela JC. Automated analysis of sleep-wake state in rats. J Neurosci Methods. 2009; 184:263–274. [PubMed: 19703489]
35. Su BL, Luo Y, Hong CY, Nagurka ML, Yen CW. Detecting slow wave sleep using a single EEG signal channel. J Neurosci Methods. 2015; 243:47–52. [PubMed: 25637866]
36. van Luijtelaar EL, Coenen AM. An EEG averaging technique for automated sleep-wake stage identification in the rat. Physiol Behav. 1984; 33:837–841. [PubMed: 6522504]
37. Vorster AP, Born J. Sleep and memory in mammals, birds and invertebrates. Neuroscience and Biobehavioral Reviews. 2015; 50:103–119. [PubMed: 25305058]
38. Yaroush R, Sullivan MJ, Ekstrand BR. Effect of sleep on memory. II. Differential effect of the first and second half of the night. J Exp Psychol. 1971; 88:361–366. [PubMed: 4326302]
39. Zhovna I, Shallom ID. Automatic detection and classification of sleep stages by multichannel EEG signal modeling. Conf Proc IEEE Eng Med Biol Soc. 2008; 2008:2665–2668. [PubMed: 19163253]

Highlights

- Method to automatically detect periods of slow wave sleep (SWS) based on intracranial EEG recordings.
- Method is based on the ratio of spectral power of slower to faster frequencies during slow wave sleep.
- Method can be performed in situations where traditional scalp EEG set-up for visual sleep staging is technically difficult.
- Method can be useful research tool when studying human memory consolidation and seizure generation in SWS.

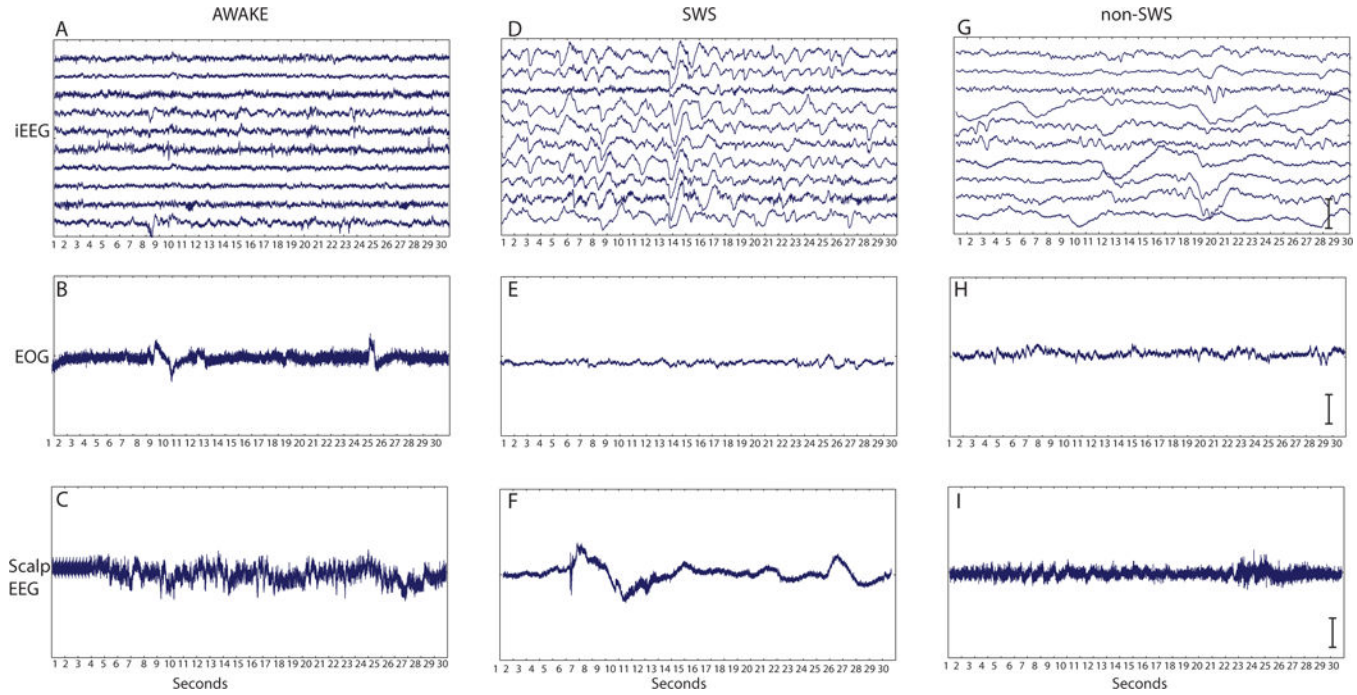

Figure 1.

Representative 30 second epoch of 10 lateral intracranial EEG recordings (Panels A, D, G), EOG (Panels B, E, H), and scalp EEG (Panels C, F, I) in patient #7 during awake (Panels A, B, C), SWS (Panels D, E, F) and non-SWS (Panels G, H, I) states. EOG, electrooculography; EEG, electroencephalography; SWS, slow wave sleep; iEEG, intracranial EEG. Vertical scale bar = 100 uV for EOG and scalp EEG and 175 uV for iEEG.

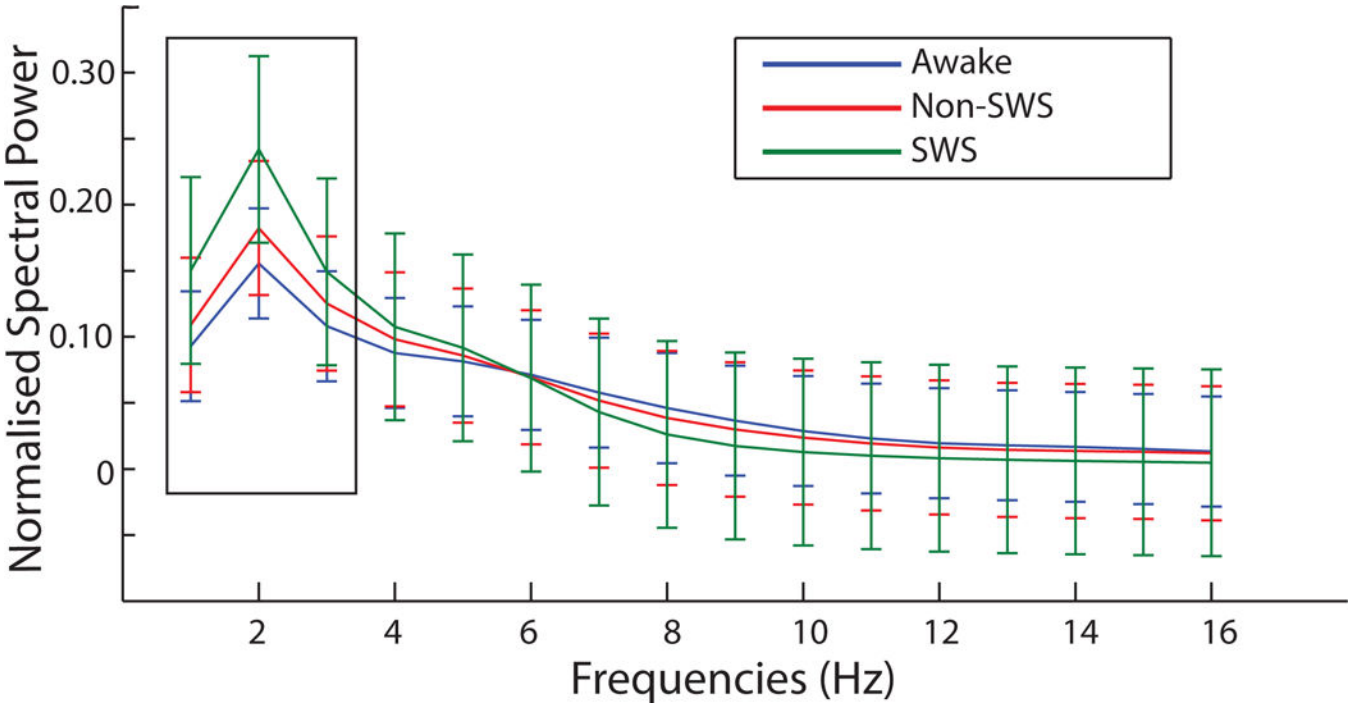

Figure 2.

Average power spectrum from the first 7 patients of intracranial EEG comparing SWS to awake states, and all other sleep stages combined. Power was normalized so that each patient contributed equally (see methods). Blue, awake; red, non-SWS; green, SWS. SWS, slow wave sleep; Results are plotted as the mean ± SD.

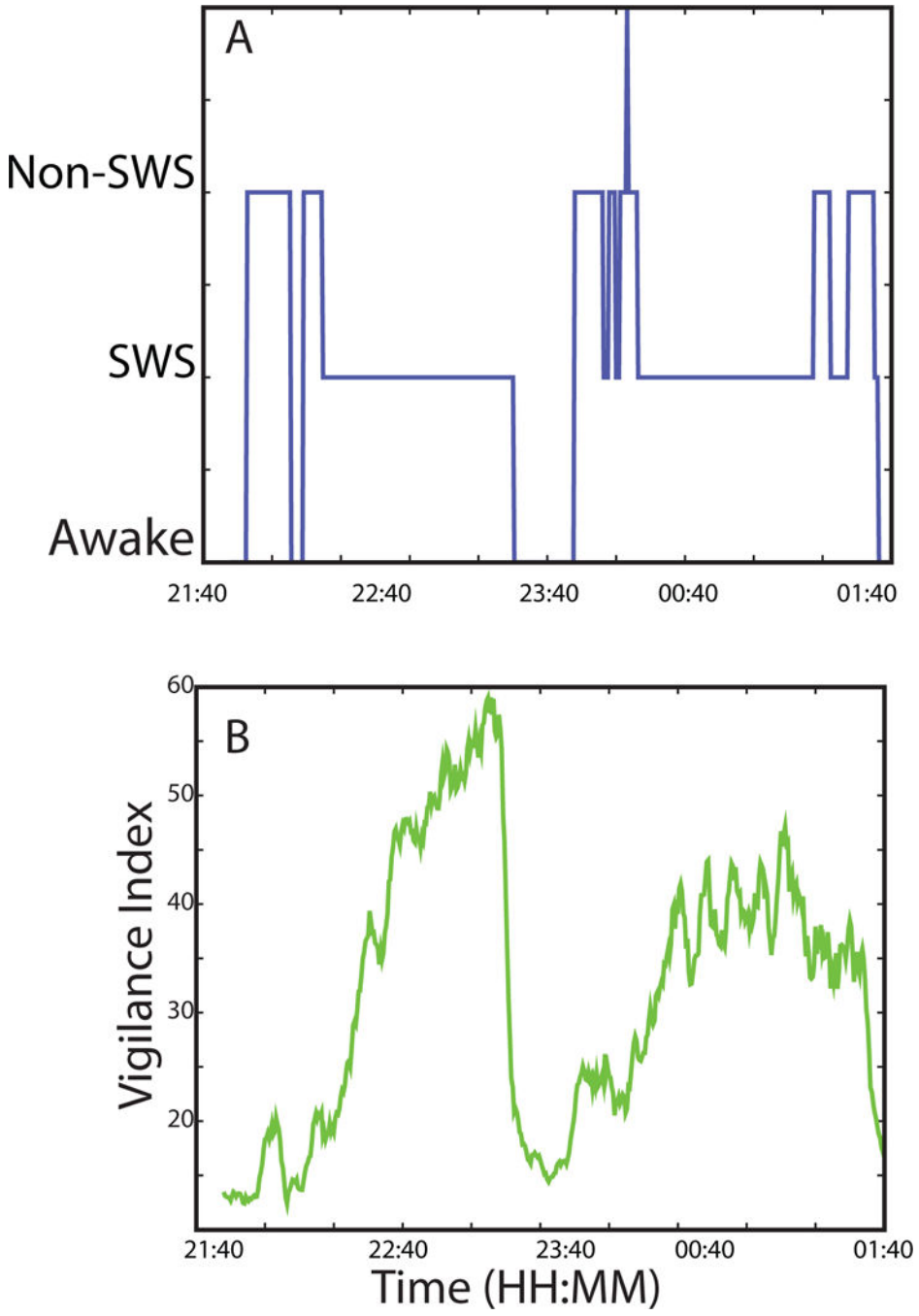

Figure 3.

Hypnogram of manual sleep staging (Blue, Panel A), and automated VI (Green Panel B) in patient #7 calculated from 9:40 pm to 1:40 am. As the patient goes into SWS, the VI index increases (Panel B). VI, vigilance index; SWS, slow wave sleep.

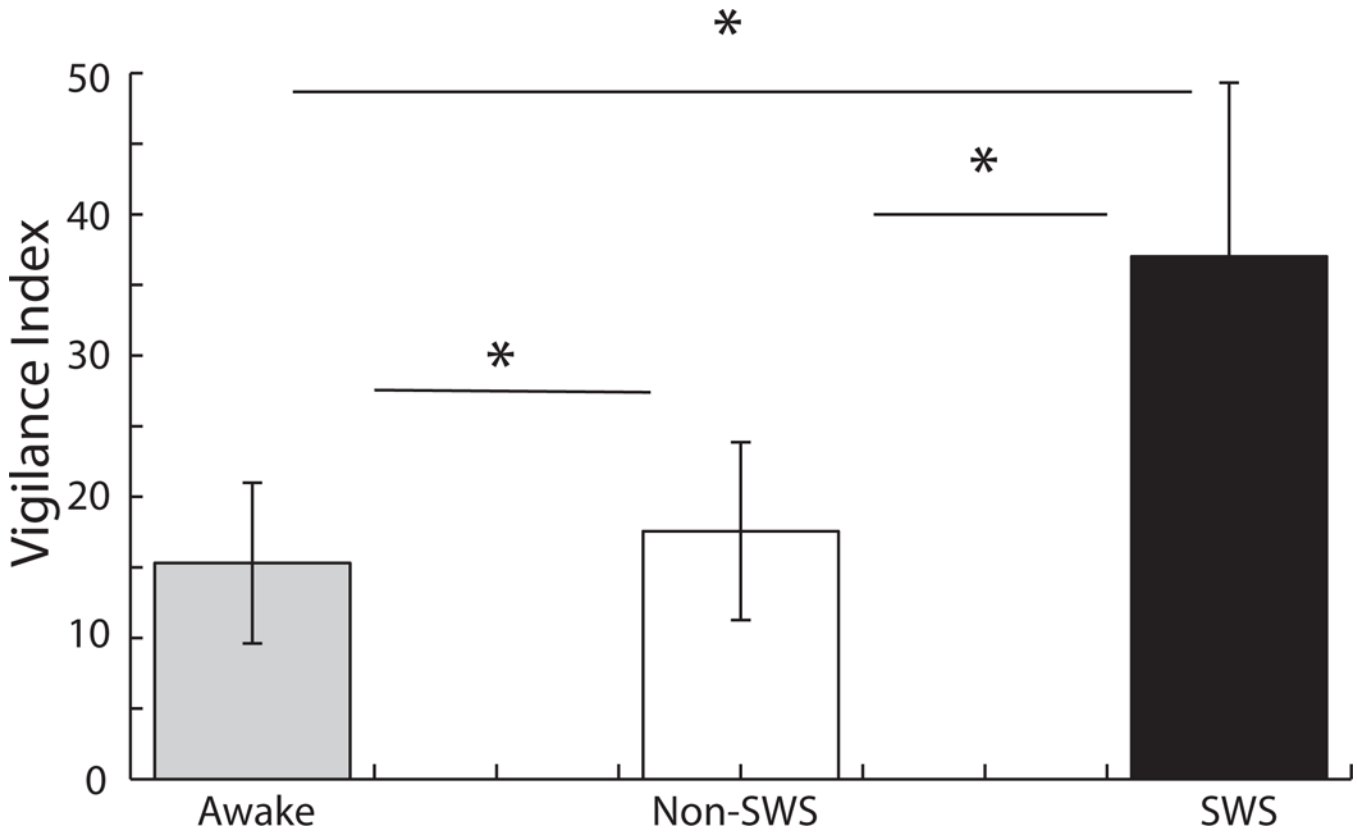

Figure 4.

The average VI for one night of sleep in patient #7 in each stage of sleep as identified by visual scoring. Black bar; SWS, grey bar; awake, white bar; non-SWS. SWS, slow wave sleep; VI, vigilance index. Results are plotted as the mean ± SD. *p<0.001.

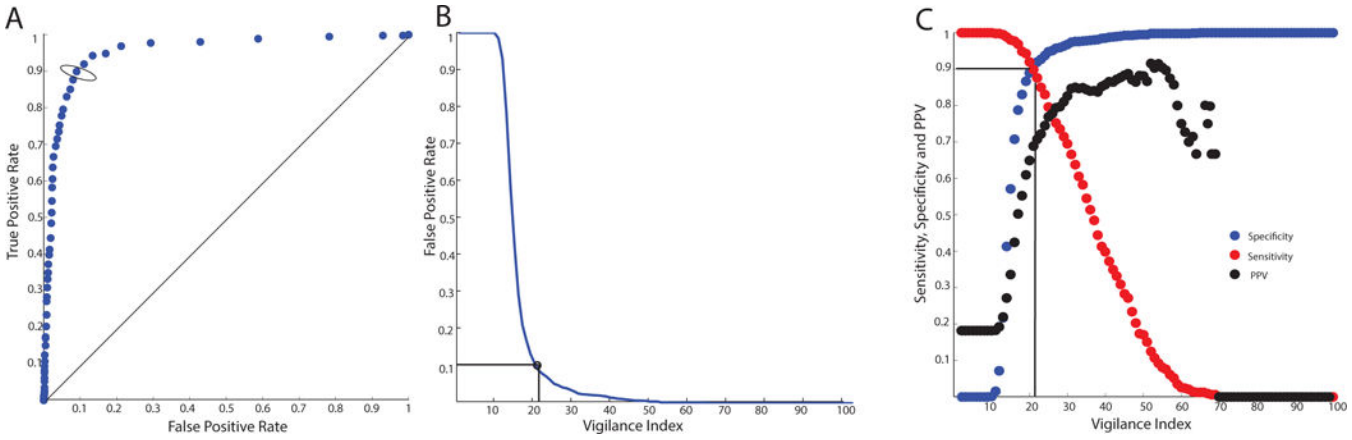

Figure 5.

Representative AUC for patient #7 for one night of automated and manual sleep staging performed by rater #1 (Panel A). The marked point represents the threshold (FP=0.1) identified by our method. The threshold of VI identified when FP is 0.1 (Panel B). The sensitivity, specificity and PPV for this threshold (Panel C). VI, vigilance index; FP, False positive rate; PPV, Positive predictive value.

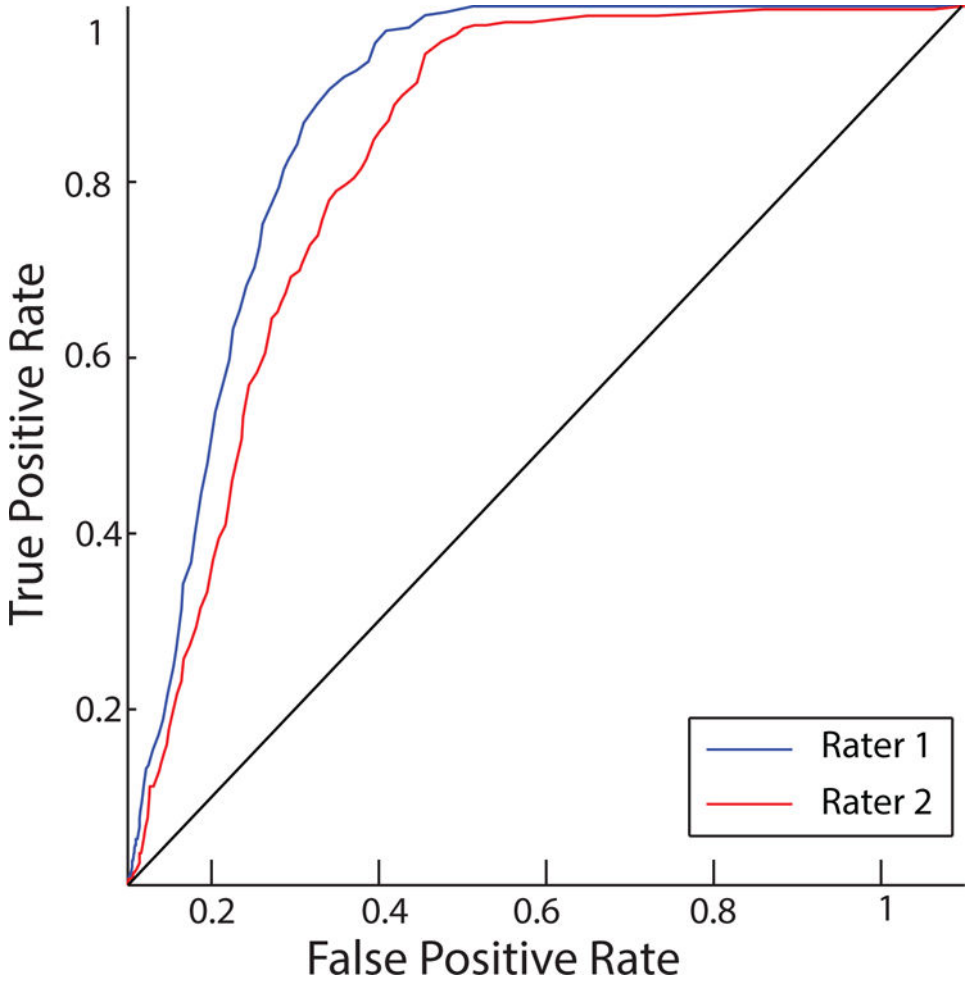

Figure 6.

Comparison of performance for automated sleep staging performed on patient #5 compared to manual scores by two separate raters. Rater 1; AUC = 0.88, κ= 0.65. Rater 2; AUC = 0.83, κ= 0.60.

Table 1

Sleep staging characteristics for each patient. AUC, Area under the curve, PPV, Positive predictive value. Patient #4 had multiple nights (marked as −N1 and −N2). Patient #5 and #8 had two visual sleep raters (marked with −R1 and −R2).

Patient ID	**1**	**2**	**3**	**4-N1**	**4-N2**	**5-R1**	**5-R2**	**6**	**7**	**8-N1**	**8-N2**	**9**	**Mean±SD**
AUC	0.96	0.96	0.92	0.96	0.96	0.88	0.83	0.89	0.95	0.89	0.89	0.81	0.91±0.05
Kappa (when FP=0.1)	0.78	0.79	0.71	0.81	0.81	0.64	0.60	0.74	0.84	0.57	0.58	0.80	0.73±0.09
Sensitivity when FP 0.1	0.90	0.90	0.80	0.90	0.86	0.8	0.75	0.80	0.90	0.81	0.80	0.80	0.83±0.05
Specificity when FP 0.1	0.85	0.85	0.80	0.90	0.85	0.8	0.75	0.80	0.90	0.81	0.80	0.80	0.83±0.04
PPV when FP 0.1 (%)	61	65	75	75	75	55	50	75	72	61	70	35	64±12.50
Kappa (threshold set for mean VI +1SD)	0.87	0.87	0.63	0.84	0.82	0.65	0.60	0.66	0.83	0.54	0.54	0.78	0.72±0.12
Mean chance kappa	0.50	0.60	0.2	0.45	0.40	0.35	0.35	0.30	0.45	0.20	0.20	0.69	0.39±0.16
Empirical p-value of Kappa (threshold set for mean VI +1SD) vs. chance	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001

Table 2

Review and comparison of characteristics of previously published automated sleep-scoring approaches using invasive electrodes.

Authors	Year	Type of analysis	Species	User-user	Auto-visual
van Luijtelaar et al.,[36]	1984	Computer based automated spectral analysis	Rat	93%	93.6%
Mamelak et al.,[21]	1988	Period-amplitude analysis using artificial neural networks (ANN)	Cat	N/A	90%
Benington et al.,[3]	1994	Computer based algorithm based on thresholding The amount of delta, theta and sigma power.	Rat	N/A	~91%
Crisler et al.,[8]	2008	107 frequency and time parameters were extracted from ECOG and EMG to train support vector machine (SVM)	Rat	86.3%	96%
Kohtoh et al.,[15]	2008	Fast Fourier transform power spectrum analysis	Rat	N/A	90%
Stephenso n et al.,[34]	2009	Automated system (ratSAS) to reject artifacts and differentiate sleep states	Rat	N/A	Wake=92% NREM 88% REM 80%.
Rytkonen et al.,[26]	2011	Open source MATLAB algorithm using naive Bayes classifier	Rats and mice	88–92%	98%